



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

## Proportion of transcriptionally active DNA virus integrants: a meta-analysis

### Citation for published version:

Marongiu, L 2017, 'Proportion of transcriptionally active DNA virus integrants: a meta-analysis', *Future virology*, vol. 12, no. 10. <https://doi.org/10.2217/fvl-2017-0063>

### Digital Object Identifier (DOI):

[10.2217/fvl-2017-0063](https://doi.org/10.2217/fvl-2017-0063)

### Link:

[Link to publication record in Edinburgh Research Explorer](#)

### Document Version:

Publisher's PDF, also known as Version of record

### Published In:

Future virology

### General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

### Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact [openaccess@ed.ac.uk](mailto:openaccess@ed.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.



# Proportion of transcriptionally active DNA virus integrants: a meta-analysis

Luigi Marongiu<sup>\*,1</sup>

<sup>1</sup>Roslin Institute, the University of Edinburgh, Easter Bush campus, EH25 9RG Edinburgh, Scotland

\* Author for correspondence: [luigi.marongiu@outlook.com](mailto:luigi.marongiu@outlook.com)

Oncoviruses are collectively responsible for over 1,000,000 new cases of cancer per year; some can integrate into the host's chromosomes. The present work was aimed at assessing the proportion of transcriptionally active viral integrants through a systematic review of the scientific publications present on the MedLine database. From the articles screened, 628 viral integrants overall were retrieved, of which 530.84 were transcriptionally active (84.53%); among the clinical samples, 264 of 323 integrants were active (81.73%). The causes for the silencing were not addressed in the articles analyzed. These findings might highlight a possible risk factor for the insurgence of cancer since some oncovirus integrants could be reactivated by stimuli of disparate nature. Further studies should address such possibility.

First draft submitted: 17 May 2017; Accepted for publication: 20 July 2017; Published online: 4 October 2017

**Keywords:** integration • latency • oncoviruses • selectable integrant • transcription silencing

About 16% of all cancer cases have been identified having an infectious aetiology [1]. Hepatitis B virus (HBV), human papilloma virus (HPV), human herpes 4 virus (HHV-4, historically known as Epstein–Barr virus [EBV]) and human herpes 8 virus (HHV-8, previously Kaposi sarcoma herpes virus) are collectively responsible for 56.2% of the incident microbial-related cancer cases, corresponding to over 1,200,000 new cases per year [2].

HBV, HPV and HHV-4 share the capability of integrating into the host's chromosomes whereas HHV-8 integrants have not been observed [3]. In addition, other DNA viruses with transformation potential can integrate: Merkel cell polyomavirus (MCPV), human herpes 6 virus (HHV-6) and, in a lesser extent, adenovirus (AdV) [4–6]. The infection with these viruses is associated with a wide range of cancer types: HBV with hepatocarcinoma; HPV with virtually all cases of cervical cancer and in a significant proportion of other tumors, for instance in the head and neck; HHV-4 with a number of malignancies, in particular Burkitt and Hodgkin lymphomas, in both immunologically deficient and competent patients; HHV-8 DNA has been retrieved in all Kaposi sarcoma lesions even in immunocompetent hosts; MCPV is recognized as the causative agent of the Merkel cell carcinoma; HHV-6 has been involved in lymphoproliferative diseases; and AdV has been shown to transform rodent cells [1,7–9].

All these viruses encode for proteins that disrupt the cell cycle and DNA damage repair response to sustain the viral replication. For example, HBV X, HPV E6/E7, HHV-8 LANA and MCPV large T proteins, all hinder the function of p53, a pivotal factor of the DNA damage response, and pRb, an oncosuppressor that prevents over-replication of the cellular DNA [10–14]. HHV-4 encodes for a plethora of early proteins with oncogenic potential; for instance, BCRF1 and BCRF1 have antiapoptotic function and the viral DNase enzyme can cause genomic instability, therefore induction of the lytic cycle can contribute to the development of malignant phenotypes [15]. Even during the lysogenic cycle, HHV-4 encodes for proteins and interfering RNAs that have oncogenic potential [16].

Integration can disrupt the transcriptional regulation of the viral oncogenes facilitating the insurgence of cancer. For example, HPV integration, which is usually accompanied by the deletion of the open reading frame (ORF) encoding the regulatory protein E2, is associated with overexpression of the oncoproteins E6 and E7 [5]. Usually integration occurs in multiple viral copies per cell; for instance, an average 36 HHV-4 integrated viral copies per cell have been identified in clinical samples [17]. A higher number of viral copies imply greater oncogene encoding capability; however, overexpression of the oncoproteins might be deleterious even for the transformed cells and an epigenetic silencing is typically established in order to avoid excessive cellular proliferation and consequent senescence [18].

Epigenetic silencing is reversible in nature and several studies have pointed out the dynamic aspects of the epigenetic control [19]. It is feasible to assume that the transcriptionally silent integrants could be reactivated under specific circumstances and they have been referred to as 'selectable' [16]. For instance, treatment of CaSki cells with the methylation inhibitor 5-azacytidine resulted in the transcriptional activation of the selectable HPV integrants [18]. Similarly, HPV expression is blocked in nontumorigenic cells expressing the transcriptional repressors MCP-1 and Jun/Fra-1 whereas it is active in the tumorigenic variants of the same cells that do not encode for these factors [20,21]. Increased calcium levels induced the exit from the latent cycle of the HHV-4-containing cell line Raji, although this line is characterized by the coexistence of both integrative and episomal forms of the virus [22]. Modifications of the cellular environment such as hypoxia and oxidative stress have been reported to reactivate transcriptionally silent HHV-4 proviruses [23,24].

Since cells containing viral integrants are typically a minority within a tissue and, by definition, transcriptionally silent viral copies do not express mRNAs or proteins, screening assays might miss the presence of the latent oncogenic proviruses in a given sample. Likewise, the immune system is not activated due to the absence of sufficient levels of viral antigens. Viral integrants have therefore the potential to remain unnoticed for long periods within a tissue. Upon reactivation, selectable integrants could accumulate oncoproteins in the affected cells, beginning the process of destabilization of the cellular biochemistry that might facilitate oncogenesis. Given this potential, selectable integrants might represent a risk factor for the insurgence of cancer and their prevalence could have been underestimated. The present study was carried out to appraise what proportion of viral integrants was actually active within the infected cells in order to assess the feasibility of such risk.

## Methods

The *MedLine* database was queried using the *PubMed* search engine with the following keywords:

((((( (((((((((human papilloma?virus\*) OR hpv?))) OR (((Epstein-Barr virus\*) OR HHV?4) OR human herpes?virus?4) OR hhv?4))) OR (((human herpes?virus?6) OR hhv?6))) OR (((((((kaposi?sarcoma?herpes?virus\*) OR kshv) OR human herpes?virus?8) OR hhv?8)))) OR (((merkel?cell?polyoma?virus\*) OR mcpv) OR mcpyv))) OR (((adeno?virus\*) OR adv))) OR (((hepatitis?b?virus) OR hbv)))) AND (((integra\*) OR pro?vir\*)) AND (((express\*) OR transcript\*))) NOT review[Publication Type].

The identified articles were imported into the *Endnote* version 7.5 [25] environment and screening of the publications was performed within this platform. Duplicates and reviews were removed and only articles in English language were selected; no limits of publication date were imposed. The articles were step-wise screened by title, abstract and content. The eligibility criterion was the presence of data related to the genetic expression of viral integrants; this was considered both in terms of viral mRNA detection or presence of viral proteins. Articles regarding the use of AdV as a gene expression system were excluded from the present analysis. The samples reported in the selected articles were also evaluated for the concomitant presence of episomal forms of the virus panel: samples indicated to contain mixtures of viral integrants and episomes were excluded from this review. In cases where multiple tests were used to assess the viral transcription, the assays providing higher transcription rate were used in order to avoid overestimation of the proportion of silent integrants. Results were plotted using *R* version 3.3.3 [26].

## Results

The *MedLine* database was queried on 7 December 2016 and 1791 articles were retrieved, with 2 duplicates. After screening, 65 articles (3.63%) were selected because they contained information related to the genetic expression from viral integrants; they all dated from 1982 onwards (Figure 1). The resulting data were stratified by sample type, virus species and detection method (Table 1). The selected 65 articles were mostly related to HPV and HBV (44.62 and 40.00%, respectively), followed by HHV-4 (7.69%), MCPV (4.62%) and HHV-6 (3.08%).

The cell lines, both primary and established, included in the present study were: 2.2.15 [27]; AM-HLH [28]; C4-I, C4-II, CaSki [29]; CU-VI-8 [30]; EC109 [31]; HeLa [29,32]; Hep3B [27,33]; hu-GK14 [34]; HX151c, HX155c, HX156c, HX160c [35]; IB-4 [36,37]; Katata [38]; ME-180 [29]; MS1 [39]; MS-B6, MS-T6 [40]; MS751 [29,41]; NAB-2 [42]; Namalwa-C18 [36]; Pal-1 [43]; PCL/PRF/5 [27]; QG-U, QG-H [44]; SiHa [29,45]; SKG-I, SKG-II, SKG-III [44]; SNU-182, SNU-354, SNU-368, SNU-387, SNU-398, SNU-423, SNU-449, SNU 475 [46]; SNU-739, SNU-761, SNU-878, SNU-886 [47]; UM-SCV-6 [48]; UD-SCC-2, UD-SCC-047 [49]; UPCI:SCC090 [50]; and UPCI:SCC154 [49].

**Table 1. Proportion of transcriptionally active viral integrants. The samples bearing viral integrants and correlated with expression data were stratified by sample type (cluster), viral species and detection methods.**

Cluster	Virus	Detection method <sup>†</sup>		Integrants		VL range <sup>‡</sup>	Ref.
		Physical status	Viral expression	Total	Actives		
Clinical samples	HBV	SB	IHC	12	2	–	[88]
		SB	NB	12	9	0.72–6	[89–92]
		SB	RT-PCR	28	19	–	[51,93,94]
	HPV	APOT	APOT	60	60	–	[95,96]
		DIPS	RT-PCR	19	16	10 <sup>–6</sup> –97	[79]
		PCR	IHC	6	1	10 <sup>–7</sup> –217	[57]
		PCR	RT-PCR	138	130	0.14–62.29	[80,97–99]
		SB	APOT	9	9	–	[52]
		SB	NB	8	2	1–100	[81,100,101]
		SB	RT-PCR	3	3	10 <sup>–5</sup> –0.5	[102]
	MCPV	DIPS	RT-PCR	2	2	10 <sup>–3</sup> –0.5	[103]
		From literature	IHC	26	11	–	[104]
	Total			323	264 (81.73%)	[34.65]	
Primary cell lines	HHV-4	FISH	FISH	1	1	–	[43]
	HBV	FISH	RT-PCR	1	1	–	[28]
		SB	NB	8	2	–	[46]
		SB	RT-PCR	4	4	–	[47]
	HPV	SB	RPA	1	1	200–300	[48]
		DIPS	RT-PCR	1	1	–	[30]
		FISH	RT-PCR	1	1	–	[50]
		SB	NB	4	4	50–2000	[35]
	HHV-6	FISH	RT-PCR	1	1	–	[38,40]
	MCPV	SB	RT-PCR	2	1	1.7–3.5	[39]
	Total			24	17 (70.83%)	[425.87]	
Established cell lines	HHV-4	FISH	FISH	2	1.84	–	[36]
		From literature	RT-PCR	1	1	–	[37]
		Sequencing	RT-PCR	1	1	1–4	[42]
	HBV	From literature	NB	3	3	–	[27]
		SB	NB	2	2	–	[33,34]
	HPV	APOT	RT-PCR	1	1	–	[31]
		DIPS	APOT	3	3	0.5–11	[49]
		From literature	NB	1	1	–	[32]
		SB	NB	13	12	–	[29,41,44,45]
	Total			27	25.84 (95.70%)	[4.13]	
Transfected cells	HHV-4	FISH	RT-PCR	1	1	–	[61]
	HBV	FISH	RT-PCR	192	192	–	[56]
		SB	EIA	5	5	–	[105]
		SB	IFA	1	1	–	[60]
		SB	NB	3	3	1–222	[106–108]
		SB	RT-PCR	2	2	–	[62]
	HPV	Sequencing	RNA-seq	1	1	–	[109]
		SB	NB	10	10	50–100	[59,110–112]
	Total			215	215 (100%)	[93.25]	
Transgenic mice	HBV	SB	NB	33	6	4–10	[53,54,58]

<sup>†</sup>APOT: Amplification of Papilloma Virus oncogene transcripts assay; DIPS: Detection of integrated papillomavirus sequences assay; EIA: Enzyme immune assay, including ELISA; FISH: Fluorescent in situ hybridization; HBV: Hepatitis B virus; HHV-4: Human herpes 4 virus; HHV-6: Human herpes 6 virus; HPV: Human papilloma virus; IFA: Immuno-fluorescence assay; IHC: Immuno-histochemistry; MCPV: Merkel cell polyomavirus; NB: Northern blot; PCR: Polymerase chain reaction assay, including quantitative methods; RIA: Radio-immuno assay; RNA-seq: mRNA deep sequencing; RPA: Ribonuclease protection assay; RT-PCR: Reverse-transcription PCR, including quantitative; SB: Southern blot; Sequencing: Both termination dye chemistry and whole genome sequencing.

<sup>‡</sup>VL: Viral load, reported in virus copies per cell; averages values per cluster are given in square brackets.

**Table 1. Proportion of transcriptionally active viral integrants. The samples bearing viral integrants and correlated with expression data were stratified by sample type (cluster), viral species and detection methods (cont.).**

Cluster	Virus	Detection method <sup>†</sup>		Integrants		VL range <sup>‡</sup>	Ref.
		Physical status	Viral expression	Total	Actives		
		SB	RIA	5	2	1–4	[55]
		SB	RT-PCR	1	1	–	[113]
	Total			39	9 (23.08%)	[4.75]	
Overall	Total			628	530.84 (84.53%)	[107.78]	

<sup>†</sup>APOT: Amplification of Papilloma Virus oncogene transcripts assay. DIPs: Detection of integrated papillomavirus sequences assay; EIA: Enzyme immune assay, including ELISA; FISH: Fluorescent in situ hybridization; HBV: Hepatitis B virus; HHV-4: Human herpes 4 virus; HHV-6: Human herpes 6 virus; HPV: Human papilloma virus; IFA: Immuno-fluorescence assay; IHC: Immuno-histochemistry; MCPV: Merkel cell polyomavirus; NB: Northern blot; PCR: Polymerase chain reaction assay, including quantitative methods; RIA: Radio-immuno assay; RNA-seq: mRNA deep sequencing; RPA: Ribonuclease protection assay; RT-PCR: Reverse-transcription PCR, including quantitative; SB: Southern blot; Sequencing: Both termination dye chemistry and whole genome sequencing.

<sup>‡</sup>VL: Viral load, reported in virus copies per cell; averages values per cluster are given in square brackets.

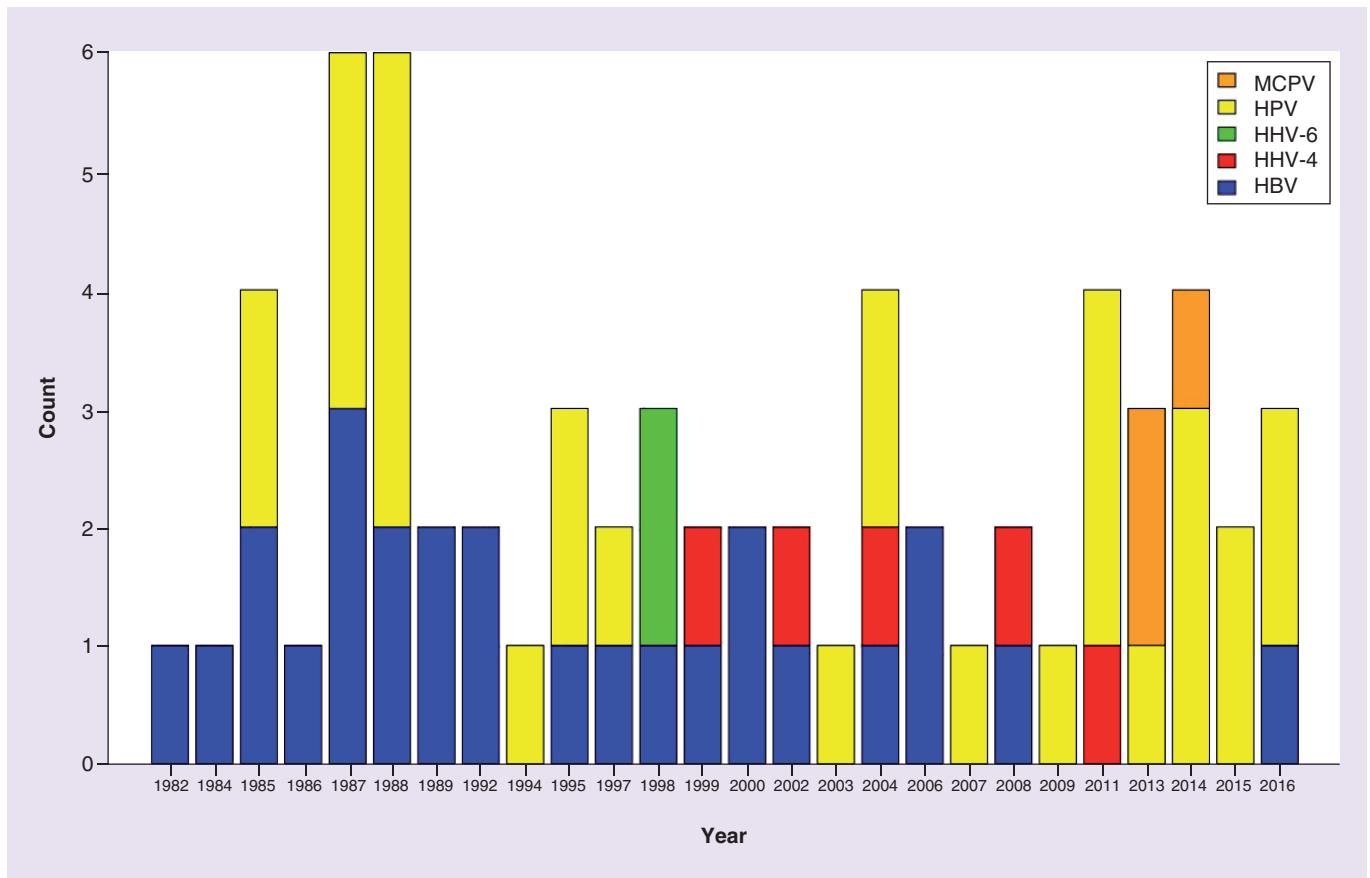
Overall the present analysis identified 628 viral integrants, of which 530.84 (84.53%) were transcriptionally active; in other words, 15.47% of the integrants were silent. One study regarding HHV-4 reported that 90–92% of the integrants in the screened cells were active [36]; the higher level was used herein to report the proportion of active viral centers, hence the fractional number. Among the clinical samples, MCPV was the least active: 46.43% of the viral integrants were shown to be transcriptionally active compared with the 57.69% of HBV and the 89.96% of HPV. Among the established cell lines, the expression rate was virtually complete (95.70%). In the primary cell line group, there were two instances where a quarter and a half of the integrants were transcriptionally active for HBV [46] and MCPV [39], respectively, whereas all the other cases were fully expressed; consequently the overall proportion of active integrants for this group was 70.83%. The transgenic mice had the lowest proportion of transcriptional activity (23.08%) in comparison to the other clusters and were infected only with HHV-4.

There were instances where the integrants were transcribing only some of the ORFs under evaluation. One study regarding HBV reported 11 out of 22 integrants as transcriptionally active (50.00%) based on the expression of the *S* gene, but only 4 (18.18%) were expressing the surface antigens (HBsAg) [51]. Transcription of the *X* gene was observed in two out of eight (25.00%) cell lines analyzed, but none expressed the *S* gene [46]. Nine out of nine (100%) HPV integrants were expressing E6 mRNAs but only seven (77.78%) were also transcribing for messenger RNAs bearing E7 [52]. Other studies reported integrants active only in certain tissues: in transgenic C57BL/6 mice, HBV particles or HBsAg were expressed in various amounts and in distinct sites in different subjects [53,54].

Some of the works gathered in the present review were the first experimental models aimed at addressing the oncogenic potential of HBV *in vivo* [53–56]. Such models recorded the vertical transmission of integrants through the germ line; remarkably, one lineage of silent integrant produced a progeny with integrants transcriptionally active in some mice and inactive in others. In the latter group, duplication of the parental viral sequence was observed and the ORF encoding the surface antigen was maintained intact. It was suggested that HBV could contain a promoter with sequences activated only in the liver and kidney due to the high amount of HBsAg often observed solely in those tissues but other results showed that, in some mice, the expression of the surface antigen could also peak in the heart and stomach. These studies also highlighted the absence of both immune response against HBsAg and signs of hepatitis in the progeny of the transfected mice; the viral antigen was postulated to be recognized as self in the mice and therefore no immune response was activated that could damage the liver. Vertical transmission in mice of integrated HBV was reported [56].

A selectable HHV-6 integrant was activated using either 12-O-tetradecanoyl-phorbol-13-acetate (TPA) or calcium ionophores [38]. This was obtained in the cell line Katata, which was derived from a Burkitt's lymphoma sample. Katata cells do not contain EBV contaminants and HHV-6 was identified as a single copy integrant (type I) located on chromosome 22. The reactivation of the virus was postulated to be dependent on protein kinase C (PKC): TPA is a direct stimulator of PKC and PKC is also activated by increased intracellular calcium levels. No infectious HHV-6 particles were produced after reactivation.

Overall the viral load (VL) of the samples gathered in the present study ranged between  $10^{-7}$  and  $10^3$  copies per cell [35,57], being highest in the primary cell lines and lowest in the established cell lines. Some studies reported either the number or the location of the integration sites, which ranged between 1 and 2 [28,33,40,41,43,45,51,58–62].



**Figure 1. Stratification by year and virus species of the scientific publications assessed in the present review.** Bar plot representing the number of articles including data regarding the expression from viral integrants ( $n = 65$ ); bars are colored according to the legend.

Remarkably, an HPV integrant with 217 copies was reported being transcriptionally silent; it was not indicated whether the HPV sequences were dispersed in more than one integration site [57].

The data regarding the integration sites were limited and did not provide information regarding a possible association between the targeted locus and transcriptional status of the provirus; neither was present a description of the chromatin context of the targeted loci. The effect of the viral integration on the expression status or phenotype of the host cell also could not be easily extrapolated. A list of the integration loci described in the retrieved articles is given in the Supplementary Table 1.

The articles selected in the present meta-analysis spanned four decades; the attempt to understand whether increases in analytical sensitivity had affected the detection of silent integrants was hampered by the heterogeneity of the data (Table 2). The proportion of transcriptionally active integrants ranged from 16.67 to 100% and overall increased steadily from 67.78% in the 1980s to 92.29% in the 2010s. In particular, the combination of Southern blot and Northern blot detected 50% of active integrants in the decade of the 1980s, 57.14% in the 1990s and 100% in the 2000s, but was not applied in the 2010s. However, even the number of samples analyzed by the Southern/Northern blot combination decreased from 74 in the 1980s to 5 in the 2000s, hence it was not possible to draw any reliable statistical conclusion from these data.

## Discussion

The present study sought to determine what proportion of oncogenic proviruses might be actually transcribing in the infected cells in order to assess whether silent copies of the integrants could be present in a given specimen. In clinical samples, such presence might constitute a risk factor for the insurgence of cancer since selectable viruses could be reactivated by several cellular stimuli at some point after the event of integration. These latent proviruses might be difficult to identify during screening.



Table 2. Detection of transcriptionally active integrants by publication period.

Period	Detection method		Integrants			Ref.
	Physical status	Viral expression	Total	Actives		
1982–1990 (67.78%)	From literature	NB	1	1	(100%)	[32]
	SB	IFA	1	1	(100%)	[60]
	SB	IHC	12	2	(16.67%)	[88]
	SB	EIA	5	5	(100%)	[105]
	SB	NB	74	37	(50.00%)	[29,35,44,45,53,54,58,89,90,92,100,101,106,108,111,112]
	SB	RIA	5	2	(40.00%)	[55]
1991–2000 (83.19%)	FISH	FISH	2	1.84	(92.00%)	[36]
	FISH	RT-PCR	2	1	(50.00%)	[38,40]
	From literature	NB	3	3	(100%)	[27]
	SB	NB	14	8	(57.14%)	[33,34,41,46,59,81,91]
	SB	RPA	1	1	(100%)	[48]
	SB	RT-PCR	10	10	(100%)	[47,62,93]
2001–2010 (85.57%)	FISH	FISH	1	1	(100%)	[43]
	APOT	APOT	3	3	(100%)	[96]
	FISH	RT-PCR	193	193	(100%)	[50,56]
	From literature	RT-PCR	1	1	(100%)	[37]
	PCR	IHC	6	1	(16.67%)	[57]
	SB	NB	5	5	(100%)	[107,110]
	SB	RT-PCR	28	19	(67.86%)	[51,94,102,113]
2011–2016 (92.29%)	Sequencing	RT-PCR	1	1	(100%)	[42]
	APOT	APOT	57	57	(100%)	[95]
	APOT	RT-PCR	1	1	(100%)	[31]
	DIPS	APOT	3	3	(100%)	[49]
	DIPS	RT-PCR	22	19	(86.36%)	[30,79,103]
	FISH	RT-PCR	2	2	(100%)	[28,61]
	From literature	IHC	26	11	(42.31%)	[104]
	PCR	RT-PCR	138	130	(94.20%)	[80,97–99]
	SB	APOT	9	9	(100%)	[52]
	SB	RT-PCR	1	1	(100%)	[39]
	Sequencing	RNA-seq	1	1	(100%)	[109]

The retrieved articles are stratified by publication year into different decades (period) and by detection method. The average proportion of transcriptionally active integrants for each decade is given in brackets.

APOT: Amplification of Papilloma virus oncogene transcripts assay; DIPS: Detection of integrated papillomavirus sequences assay; EIA: Enzyme immune assay, including ELISA; FISH: Fluorescent *in situ* hybridization; IFA: Immuno-fluorescence assay; IHC: Immuno-histochemistry; NB: Northern blot; PCR: Polymerase chain reaction assay, including quantitative methods; RIA: Radio-immuno assay; RNA-seq: mRNA deep sequencing; RPA: Ribonuclease protection assay; RT-PCR: Reverse-transcription PCR, including quantitative; SB: Southern blot; Sequencing: Both termination dye chemistry and whole genome sequencing.

The information regarding the proportion of actively transcribing integrants was obtained by a meta-analysis of scientific publications retrieved from the MedLine database. The results gathered in the present study showed that overall about 85% of the integrants were expressing viral genes; in the clinical samples, the ratio was even lower (~82%). The data were based mostly on HPV and HBV; the lack of publications regarding HHV-8 and AdV might confirm the absence of integration of these viruses, particularly in clinical samples. The seminal work carried out by Doerfler and collaborators has highlighted the consequences of AdV type 12 integration on the viral and cellular transcription following the epigenetic response of the cell to integration [63,64]. However, in more recent years, this line of research has been replaced by a more heterogeneous use of AdV type 12 as a gene delivery system;

these articles were not included in the present meta-analysis because they dealt with genes not originally present in AdV.

One article collected herein reported a silent HHV-6 integrant becoming transcriptionally active after modifying the cellular environment [38], supporting the case for a delayed viral reactivation. Reactivation of HHV-6 integrants has also been reported during coinfection with *Chlamydia trachomatis* and in immunocompromised patients [65,66]. Although no infectious particles were produced in both cases, possibly due to loss of information in the flanking repeated sequences that might affect viral assembly, excision of the viral integrants with generation of circular episomes with transcriptional capability was observed. Restoration of transcription from integrated and excised HHV-6 genomes has been linked to pathologies such as angina pectoris and it has been suggested that the process of integration and excision could facilitate the insurgence of senescence and chromatin instability in the affected cells [67]. Even in absence of viral particles, reactivation of HHV-6 integrants can be considered a risk factor for the insurgence of several pathologies including cancer. The reactivation of HHV-6 in the mentioned paper was obtained by modifying the cellular calcium level exposing the cell to either TPA or ionophores; ultimately, PKC was regarded as involved in the termination of the silencing pathway [56]. PKC has been also involved in the reactivation of latent EBV in the cell line Raji [22].

Latency could facilitate the dissemination of the HHV-6 integrants and eventually lead to their fixation in the human population. In the case of human immunodeficiency virus (HIV), the evolutionary advantage of transcriptional silencing is not immediately evident: latency is an inherent biochemical activity of HIV that pre-dates the advent of antiretroviral therapies, thus it must have an advantage in the natural history of the infection; however the gain is not apparent for a virus whose basic reproductive strategy is based on progeny with high VLs [68]. HIV has been suggested to be capable of surviving adverse physiological conditions and inefficient transmission by shielding itself in a small subset of lymphocytes as inactive provirus [68]. Similarly, HHV-6 could remain hidden in the cells, diffuse vertically and become activated when the cellular environment is modified. However horizontal transmission is expected to be precluded for the integrants since the excited viral genomes will lack information at the genome's extremities with respect to the parental episomal virus. HHV-6 integration is due to regions of homology to the human telomeres in the viral terminal repeats that are also present in human herpes 7 virus (HHV-7) and Marek disease virus (MDV) [69–71]; however, HHV-7 is not found integrated whereas HHV-6 and MDV are retrieved as integrants [72]. Further work is required to determine the molecular differences that set apart HHV-7 from HHV-6 and MDV.

The present meta-analysis identified publications regarding mice models that reported the transmission of integrated HBV from spermatozoa to the oocyte with generation of carrier embryos [53–56]. Vertical transmission of HBV is a recognized route of dissemination of HBV, however this modality normally involves infectious particles [73]. These aforementioned studies reported *in vitro* models where integration was obtained by transfecting a plasmid into spermatozoa; nevertheless, integrated HBV genomes have been retrieved in the semen of human patients [74]. Since integrated HHV-6 can also be vertically transmitted via germ line [5], it can be speculated that integrated oncoviruses could be disseminated not only by infection and their prevalence could be wider than currently assumed.

The data collected herein were disparate in terms of viral species, samples and detection method; stratification into clusters was carried out in order to define patterns that could help understand the viral behavior. These clusters had different bias toward viral integration: transgenic mice and transfected cells are actively screened for the presence of active integrants; clones of cell within a line are selected due to their replication rate and the viral expression might provide an advantage *in vitro*. Henceforth, it was not surprising that the cell lines, particularly the transfected ones, had the highest proportion of active integrants among the clusters analyzed herein. Clinical samples and in a lesser extent, primary cell lines provided a better understanding of the viral biology in physiological conditions. However the samples gathered in the present study varied in terms of gender, age, ethnicity, length of infection, previous pathologies and presence of other risk factors. Nevertheless, these disparities provided a snapshot of the viral biochemistry under several conditions and the results suggested that each of the oncoviruses included in the present work had the potential to generate silent integrants.

In addition, the analytical tools were diverse and with a different bias toward detection of active integrants. For instance, the amplification of papillomavirus oncogene transcripts (APOT) assay specifically amplifies viral messengers from transcriptionally active integrants but cannot detect the integrants directly [52], therefore silent integrants are missed. On the other end, detection of integrated papillomavirus sequences (DIPS) is used to characterize the sequence structure of the viral integrants, particularly at the virus–cell junctions, but it is not



designed to query the transcriptional activity [75], thus it cannot assess the expression status of the proviruses. The combination of APOT and DIPS was applied only once in the studies collected herein [49]. The nucleic acid amplification and blotting methods act on populations of cells and do not have the resolution to assess the transcriptional activity of individual integrants. Fluorescent *in situ* hybridization (FISH) has single cell resolution but has limited throughput therefore it is likely to miss rare integrants during screening procedures. Nonetheless, VL was never assessed by FISH in the scientific publications gathered in the present study, suggesting that despite its single-cell resolution, this technique is not suitable for quantitative purposes.

Moreover, the techniques were applied in different decades and a bias toward increased analytical sensitivity must be considered. The average proportion of active integrants increased with the publication year; such result could be an artefact of the clustering method used to report the data. Further work is required to assess whether the analytical sensitivity had changed the detection of viral integrants and their transcriptional status in recent years.

Most of the retrieved articles did not report VL data, thus it was not possible to assess whether the oncogene expression was derived from a single viral center or from a pool of proviruses. The VL from the retrieved scientific articles spanned ten orders of magnitude; despite such wide range, some observations could be made. Very low VLs could simply imply that few cells containing high amounts of virus were present in a background of neutral cells; the actual VL in these ‘jackpot’ cells [76] was unknown. These potential reservoir cells with high VL might be present in a given tissue at low concentration; should the number of noninfected cells be great enough, it is feasible to assume that their presence could be masked and their activity remain undetected. In cell lines, neutral cells are typically depleted, therefore the VL is expected to be higher than in clinical samples; counter-intuitively, though, the VL of the established cell lines was lower than that of the primary cells and clinical samples. It might be postulated that in the established lines, cells with few active proviruses were selected in order to provide the right amount of viral proteins to stimulate the cell replication without triggering senescence, as previously suggested [18]. In primary cell lines and clinical samples, such selection could not have been fully unfolded, explaining the higher VL. Among the articles retrieved in the present review, there were no follow-up studies that could assess the dynamics of VL along the passaging of primary cell lines and their evolution toward established cell lines, thus this interpretation could not be tested.

The identification of a silent integrant does not imply the presence of a selectable integrant. The recombination mechanism that is involved in the integration phase is usually accompanied by loss of sequences in both the host's chromosomes and in the viral genome [77]. As a consequence, the integrant might lose its promoter region and become transcriptionally incompetent. Tandem integrants (type II) such as head and tail concatamers are highly likely to contain intact, thus transcriptionally competent, viral sequences while having disruptions only at the extremities of the array. In addition, a recent work has shown that tandem HPV integrants can behave as enhancers for the transcription of the viral oncogenes [78], thus type II integrants are expected to have transcriptional capability and high expression rate. On the other hand, transcription from individual integrants (type I) might be assumed to be restricted by the functional integrity of the viral sequence, although host's promoters might compensate for the loss of the viral ones. There was no direct data in the present work that could corroborate such assumption: the scientific publications recording clinical samples with high viral copy number only reported the VL data for the overall group rather than for each sample, therefore it was not possible to assess whether there was a relation between viral copy numbers and transcriptional status [79–81]. Furthermore, it has been suggested that type II pro-viruses might undergo epigenetic silencing [19]; this process can overcome the physical status of the integrants and disrupt an association between VL and transcriptional activity.

Ultimately, the actual expression profile of a given integrant is due to its chromatin conformation [82], which can be altered by epigenetic mechanisms such as the methylation status of the viral sequences; however the studies collected herein did not include an epigenetic analysis. It would have been interesting, for instance, to assess whether the transcriptionally inactive integrant with a VL of  $10^2$  identified in a clinical samples [57] was silent because of epigenetic suppression or for other reasons such as loss of the viral promoter. Remarkably, transgenic mice showed the lowest proportion of actively transcribing proviruses, clearly indicating that despite bearing at least one copy of integrated virus in each cell, the viral genetic expression was modulated by the cellular environment; epigenetic mechanisms were evoked to explain the transcriptional silencing [53,54].

It is logical to assume that only a subset of the transcriptionally silent integrants identified in the present study were truly selectable in nature. Nevertheless, given the high prevalence of infection of the viruses included in this review [83] and the non-negligible proportion of silent integrants identified herein, which ranged from one in five to one in ten of all integrants, it is possible that a significant number of people could carry selectable integrants.

These viruses could be expected to be difficult to identify by routine diagnostic and reactivate years after the actual infection, so that the clinical features that could facilitate the diagnosis might be forgotten.

Based on these data, two speculative scenarios can be drawn. In the first one, a subject becomes infected horizontally via the typical routes of transmission of the oncoviruses; integration might occur in a small subset of cells within a given tissue. The provirus could remain occult for a certain amount of time until a cellular stimulus would trigger an environmental modification, reactivating the expression of viral oncogenes, facilitating the development of the malignant phenotype as a result. Latent rare integrants are expected to be a minority within a tissue, so their presence would likely be missed by screening tests. For instance, a limit of detection of one genome equivalent per microliter of sample and a minimum of 20% integration were reported for assays assessing VL and physical status of HPV, respectively [84]; for HBV testing, the recommended VL cutoff is in the range of 1–1000 genome equivalents per microliter of samples [85]. Specimens containing very low amounts of viral genomes that are formed by a limited number of cells with high VL among a high number of neutral cells or where few integrants are coexisting with many native episomal forms of the virus would be likely to give a false negative result during screening testing. In addition, since the reactivation would appear a long time after the infection, it is plausible to assume that no specific tests targeting oncoviruses would be implemented and the presence of the integrants might remain undetected. The immune response, in particular the action of interferon, is able to deplete viral episomes from the infected cells [86]; however integrants cannot be removed by this mechanism and can therefore persist after viral clearance. Interestingly, it has been reported that interferon can facilitate viral integration [87].

The second scenario is more difficult to define and it is based on the vertical transmission of oncoviruses via germ line. This is a recognized route of infection of HHV-6, which is characterized by abnormal high VL in serum samples and that has resulted in mistreatment of patients [8]. The outcome of this modality of infection can be outlined using the HBV mice models [54–56]. Should viral sequences be present in every cell of a subject rather than within a single tissue, it should be relatively easy to identify the presence of the oncoviruses by molecular biology assays. On the other hand, the presence of viral antigens that could be produced in low quantities in fetuses and newborns would be recognized as self by the immune system; consequently detection of the virus by serology assays might be affected. Probably the reactivation of these constitutive proviruses would occur only in specific tissues rather than being disseminated, in which case the two scenarios will essentially converge. In both scenarios, selectable integrants might facilitate the insurgence of the malignant phenotype and therefore represent a risk factor for the development of cancer.

## Conclusion

The present study was aimed at assessing what proportion of oncogenic viral integrants was actively transcribing; this task was addressed by a meta-analysis of the literature present on the MedLine database. The data showed that about 85% of the integrants were actively expressing viral products. The converse implied that about one in five to one in ten of the integrants could be in a state of transcriptional latency. The data collected could not explain the reasons for the transcriptional inactivity of these proviruses. The presence of a non-negligible proportion of oncogenic viruses that could integrate and remain transcriptionally silent until reactivated by a cellular stimulus might represent a risk factor for the insurgence of several types of cancer. Further studies are needed to assess whether such risk has a statistical and clinical foundation.

## Future perspective

The data collected in the present review highlighted the possibility of the presence of selectable oncovirus integrants in the general population. The vast majority of the articles collected herein were focused on HBV and HPV. More studies related to the other oncoviruses with integrative capabilities might elucidate the mechanism of integration and assess whether a common pathway is present. For instance, one answer that future work could address is why these viruses show such differences in the frequency of integration. Both HHV-4 and HHV-8, like the other members of the *Herpesviridae* family, have a linear genome that could, in theory, lead to recombination; nevertheless, the former has been reported as a provirus only in cell lines whereas the latter has not been identified in the integrative form at all. HBV, HPV and MCPV possess circular DNA genomes and such feature could explain why these viruses have a higher frequency of integration than the herpesviruses. This is a somewhat counter-intuitive concept: first, circular DNA is less prone to integration, hence HBV, HPV and MCPV should integrate less than the herpesviruses; second, the genomes of the herpesviruses are recircularized in the nucleus of the infected cells, thus they should undergo the same mechanisms and frequencies of integration of HBV, HPV and MCPV. HHV-6

is a peculiar virus that, due to the sequences with telomeric homology present in its genome follows a separate pattern of integration than the other viruses included in the present review.

In recent times, the combination of whole genome sequencing and RNA-seq analysis has become more and more common. Such approach provides valuable description of the viral integrants in terms of their mutations, loss of coding sequences and acquisition of cellular promoters' control over transcription; the mRNA data complement the picture indicating which integrants are actively transcribing. It can be anticipated that in the forthcoming years, more emphasis will be given to longitudinal studies that will focus on the nontranscribing integrants to understand whether they might be activated upon given cellular stimuli. Methylation status and chromatin conformation will also provide useful information regarding the mechanisms of latency.

More studies are required to assess the true prevalence of latent oncovirus integrants; given the expected rarity of the integration events and the absence of viral transcription, the screening methods should be aimed at analyzing high number of cells within a specimen and should have high sensitivity. PCR methods are characterized by high analytical sensitivity but should be coupled by selection methods that could reduce the bias introduced by noninfected cells. Single-cell PCR is a possible approach for such screening; flow-FISH is another possible way, the latter taking advantage of the high-throughput of the flow cytometer assay with the single cell resolution of the FISH assay.

#### Acknowledgements

The author gives special thanks to MD Curran (Public Health England, Clinical Microbiology and Public Health Laboratory, Addenbrooke's Hospital, Cambridge, England) for his intellectual support.

#### Financial & competing interests disclosure

This work was supported by the Biotechnology and Biological Sciences Research Council (BB/J004324/1; BBS/E/D/20241864). The author has no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

#### Supplementary data

To view the supplementary data that accompany this paper please visit the journal website at: [www.futuremedicine.com/doi/full/10.2217/fvl-2017-0063](http://www.futuremedicine.com/doi/full/10.2217/fvl-2017-0063)

#### Summary points

- DNA viruses with oncogenic potential have been reported to be integrated in cell lines and clinical samples. Not all oncovirus integrants are transcriptionally active.
- Although in some cases transcriptional inactivity could simply be due to rearrangement of the viral sequences that resulted in the loss of regulatory or oncogene sequences, in other cases a true latency might have been established. Transcriptionally competent virus integrants in latent state have been defined as 'selectable'.
- Transcriptional silencing is obtained through epigenetic mechanisms, which are reversible. Several factors can trigger the modification of the epigenetic pressure on the viral integrants inducing reactivation of the expression of oncogenes, facilitating the insurgence of cancer.
- The present review of scientific publications present in the MedLinedatabase assessed that 84.53% integrants were actively transcribing for viral products; in clinical samples alone, the proportion of transcriptionally active integrants fell to 81.73%. The results implied that the remainder of the integrants were either in a status of latency or incapacitated to transcribe, although the articles analyzed did not address the reasons of the transcriptional silencing.
- Given the high prevalence of the oncoviruses infections, even if a fraction of the silent viral copies observed herein represented genuine selectable integrants, it can be speculated that a non-negligible proportion of the population might carry viral oncogenes that could be reactivated by a plethora of cellular stimuli. Therefore, these selectable viral integrants might represent a potential risk factor in oncogenesis.
- Further studies are required to assess whether this potential risk is concrete.

#### References

- 1 de Martel C, Ferlay J, Franceschi S *et al.* Global burden of cancers attributable to infections in 2008: a review and synthetic analysis. *Lancet Oncol.* 13, 607–615 (2012).

- 2 Plummer M, De Martel C, Vignat J, Ferlay J, Bray F, Franceschi S. Global burden of cancers attributable to infections in 2012: a synthetic analysis. *Lancet Oncol.* 4, e609–e616 (2016).
- 3 Ahuja R, Jamal A, Nosrati N *et al.* Human oncogenic viruses and cancer. *Curr. Sci.* 107(5), 768–785 (2014).
- 4 Haugg A, Rennspies D, zur Hausen A *et al.* Fluorescence *in situ* hybridization and qPCR to detect Merkel cell polyomavirus physical status and load in Merkel cell carcinomas. *Int. J. Cancer* 135, 2804–2815 (2014).
- 5 Pellett PE, Ablashi DV, Ambros PF *et al.* Chromosomally integrated human herpesvirus 6: questions and answers. *Rev. Med. Virol.* 22(3), 144–155 (2012).
- 6 Janovitz T, Klein IA, Oliveira T *et al.* High-throughput sequencing reveals principles of adeno-associated virus serotype 2 integration. *J. Virol.* 87(15), 8559–8568 (2013).
- 7 Spurgeon ME, Lambert PF. Merkel cell polyomavirus: a newly discovered human virus with oncogenic potential. *Virology* 435(1), 118–130 (2013).
- 8 Agut H, Bonnafous P, Gautheret-Dejean A. Laboratory and clinical aspects of human herpesvirus 6 infections. *Clin. Microbiol. Rev.* 28(2), 313–335 (2015).
- 9 Speiseder T, Hormann-Sieber H, Rodríguez E, Dobner T. Efficient transformation of primary human mesenchymal stromal cells by Adenovirus early region 1 oncogenes. *J. Virol.* 91(1), e01782–16 (2017).
- 10 Chan C, Thurnherr T, Wang J *et al.* Global rewiring of p53 transcription regulation by the hepatitis B virus X protein. *Mol. Oncol.* 10(8), 1–13 (2016).
- 11 Tommasino M. The human papillomavirus family and its role in carcinogenesis. *Semin. Cancer Biol.* 26, 13–21 (2014).
- 12 Borchert S, Czech-Sioli M, Neumann F *et al.* High-affinity Rb binding, p53 inhibition, subcellular localization and transformation by wild-type or tumor-derived shortened Merkel cell polyomavirus large T antigens. *J. Virol.* 88(6), 3144–3160 (2014).
- 13 Schulz TF, Cesarman E. Kaposi sarcoma-associated herpesvirus: mechanisms of oncogenesis. *Curr. Opin. Virol.* 14, 116–128 (2015).
- 14 Jung JK, Arora P, Pagano JS, Jang KL. Expression of DNA methyltransferase 1 is activated by hepatitis B virus X protein via a regulatory circuit involving the p16INK4a-cyclin D1-CDK 4/6-pRb-E2F1 pathway. *Cancer Res.* 67(12), 5771–5778 (2007).
- 15 Li H, Liu S, Hu J *et al.* Epstein–Barr virus lytic reactivation regulation and its pathogenic role in carcinogenesis. *Int. J. Biol. Sci.* 12(11), 1309–1318 (2016).
- 16 Giudice A, D’Arena G, Crispo A *et al.* Role of viral miRNAs and epigenetic modifications in Epstein–Barr virus-associated gastric carcinogenesis. *Oxid. Med. Cell. Longev.* 2016 (2016).
- 17 Batinica M, Akgul B, Silling S, Mauch C, Zigrino P. Correlation of Merkel cell polyomavirus positivity with PDGFR- $\alpha$  mutations and survivin expression in Merkel cell carcinoma. *J. Dermatol. Sci.* 79(1), 43–49 (2015).
- 18 Van Tine B, Kappes J, Banerjee N. Clonal selection for transcriptionally active viral oncogenes during progression to cancer. *J. Virol.* 78(20), 11172–11186 (2004).
- 19 Johannsen E, Lambert PF. Epigenetics of human papillomaviruses. *Virology* 445(1–2), 205–212 (2013).
- 20 Rösl F, Lengert M, Albrecht J *et al.* Differential regulation of the *JE* gene encoding the monocyte chemoattractant protein (MCP-1) in cervical carcinoma cells and derived hybrids. *J. Virol.* 68(4), 2142–2150 (1994).
- 21 Soto U, Das BC, Lengert M, Finzer P, zur Hausen H, Rösl F. Conversion of HPV-18-positive nontumorigenic HeLa-fibroblast hybrids to invasive growth involves loss of TNF- $\alpha$  mediated repression of viral transcription and modification of the AP-1 transcription complex. *Oncogene* 18(21), 3187–3198 (1999).
- 22 Faggioni A, Zompetta C, Grimaldi S, Barile G, Frati L, Lazdins J. Calcium modulation activates Epstein–Barr virus genome in latently infected cells. *Science* 232(4757), 1554–1556 (1986).
- 23 Jiang JH, Wang N, Li A *et al.* Hypoxia can contribute to the induction of the Epstein–Barr virus (EBV) lytic cycle. *J. Clin. Virol.* 37(2), 98–103 (2006).
- 24 Lassoued S, Gargouri B, El Feki AEF, Attia H, Van Pelt J. Transcription of the Epstein–Barr virus lytic cycle activator BZLF-1 during oxidative stress induction. *Biol. Trace Elem. Res.* 137(1), 13–22 (2010).
- 25 Eapen BR. EndNote 7.0. *Indian J. Dermatol. Venerol. Leprol.* 72(2), 165–166 (2006).
- 26 R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria.
- 27 Farshid M, Tabor E. Expression of oncogenes and tumor suppressor genes in human hepatocellular carcinoma and hepatoblastoma cell lines. *J. Med. Virol.* 38(4), 235–239 (1992).
- 28 Hayashida M, Daibata M, Tagami E *et al.* Establishment and characterization of a novel Hodgkin lymphoma cell line, AM-HLH, carrying the Epstein–Barr virus genome integrated into the host chromosome. *Hematol. Oncol.* (September), 1–9 (2016).
- 29 Yee C, Krishnan-Hewlett I, Baker C, Schlegel R, Howley P. Presence and expression of human papillomavirus sequences in human cervical carcinoma cell lines. *Am. J. Pathol.* 119(3), 361–366 (1985).
- 30 Bryant D, Onions T, Raybould R *et al.* mRNA sequencing of novel cell lines from human papillomavirus type-16 related vulval intraepithelial neoplasia: consequences of expression of HPV16 E4 and E5. *J. Med. Virol.* 86, 1534–1541 (2014).

- 31 Zhang K, Li JT, Li SY, Zhu LH, Zhou L, Zeng Y. Integration of human papillomavirus 18 DNA in esophageal carcinoma 109 cells. *World J. Gastroenterol.* 17(37), 4242–4246 (2011).
- 32 Inagaki Y, Tsunokawa Y, Takebe N *et al.* Nucleotide sequences of cDNAs for human papillomavirus type 18 transcripts in HeLa cells. *J. Virol.* 62(5), 1640–1646 (1988).
- 33 Su TS, Hwang WL, Yauk YK. Characterization of hepatitis B virus integrant that results in chromosomal rearrangement. *DNA Cell Biol.* 17(5), 415–425 (1998).
- 34 Nakamichi N, Noda A, Yonezu T, Koike K, Matsumura T. Structure and expression of integrated hepatitis B virus genes in an HBs antigen producing human cell line (huGK-14). *Cytotechnology* 25(1–3), 61–70 (1997).
- 35 Spence RP, Murray A, Banks L, Kelland LR, Crawford L. Analysis of human papillomavirus sequences in cell lines recently derived from cervical cancers. *Cancer Res.* 48(2), 324–328 (1988).
- 36 Szeles A, Falk KI, Imreh S, Klein G. Visualization of alternative Epstein–Barr virus expression programs by fluorescent *in situ* hybridization at the cell level. *J. Virol.* 73(6), 5064–5069 (1999).
- 37 Fiorini S, Ooka T. Secretion of Epstein–Barr virus-encoded BARF1 oncoprotein from latently infected B cells. *Virol. J.* 5(1), 70 (2008).
- 38 Daibata M, Taguchi T, Taguchi H, Miyoshi I. Integration of human herpesvirus 6 in a Burkitt's lymphoma cell line. *Br. J. Haematol.* 102(5), 1307–1313 (1998).
- 39 Guastafierro A, Feng H, Thant M *et al.* Characterization of an early passage Merkel cell polyomavirus-positive Merkel cell carcinoma cell line, MS-1 and its growth in NOD scid gamma mice. *J. Virol. Methods* 187(1), 6–14 (2013).
- 40 Daibata M, Taguchi T, Kubonishi I, Taguchi H, Miyoshi I. Lymphoblastoid cell lines with integrated human herpesvirus type 6. *J. Hum. Virol.* 1(7), 475–481 (1998).
- 41 Geisbill J, Osmer U, Dürst M. Detection and characterization of human papillomavirus type 45 DNA in the cervical carcinoma cell line MS751. *J. Gen. Virol.* 78(Pt 3), 655–658 (1997).
- 42 Luo W-J, Takakuwa T, Ham MF *et al.* Epstein–Barr virus is integrated between REL and BCL-11A in American Burkitt lymphoma cell line (NAB-2). *Lab Invest.* 84(9), 1193–1199 (2004).
- 43 Daibata M, Taguchi T, Nemoto Y *et al.* Epstein–Barr virus (EBV)-positive pyothorax-associated lymphoma (PAL): chromosomal integration of EBV in a novel CD2-positive PAL B-cell line. *Br. J. Haematol.* 117(3), 546–557 (2002).
- 44 Shirasawa H, Tomita Y, Sekiya S. Integration and transcription of human papillomavirus type 16 and 18 sequences in cell lines derived from cervical carcinomas. *J. Gen. Virol.* 68(2), 583–591 (1987).
- 45 el Awady MK, Kaplan JB, O'Brien SJ, Burk RD. Molecular analysis of integrated human papillomavirus 16 sequences in the cervical cancer cell line SiHa. *Virology* 159(2), 389–398 (1987).
- 46 Park JG, Lee JH, Kang MS *et al.* Characterization of cell lines established from human hepatocellular carcinoma. *Int. J. Cancer* 62(3), 276–282 (1995).
- 47 Lee J-H, Ku J-L, Park YJ, Lee K-U, Kim W-H, Park J-G. Establishment and characterization of four human hepatocellular carcinoma cell lines containing hepatitis B virus DNA. *World J. Gastroenterol.* 5(4), 289–295 (1999).
- 48 Hietanen S, Grenman S, Syrjänen K *et al.* Human papillomavirus in vulvar and vaginal carcinoma cell lines. *Br. J. Cancer* 72, 134–139 (1995).
- 49 Olthof NC, Huebbers CU, Kolligs J *et al.* Viral load, gene expression and mapping of viral integration sites in HPV16-associated HNSCC cell lines. *Int. J. Cancer* 136(5), E207–E218 (2015).
- 50 Ragin CCR, Reshmi SC, Gollin SM. Mapping and analysis of HPV16 integration sites in a head and neck cancer cell line. *Int. J. Cancer* 110(5), 701–709 (2004).
- 51 Wang Y, Wu MC, Sham JS *et al.* Different expression of hepatitis B surface antigen between hepatocellular carcinoma and its surrounding liver tissue, studied using a tissue microarray. *J. Pathol.* 197(5), 610–616 (2002).
- 52 Vojtechova Z, Sabol I, Salakova M *et al.* Analysis of the integration of human papillomaviruses in head and neck tumors in relation to patients' prognosis. *Int. J. Cancer* 138(2), 386–395 (2016).
- 53 Araki K, Miyazaki J, Hino O *et al.* Expression and replication of hepatitis B virus genome in transgenic mice. *Proc. Natl Acad. Sci. USA* 86(1), 207–211 (1989).
- 54 Chisari F, Pinkert C, Milich D *et al.* A transgenic mouse model of the chronic hepatitis B surface antigen carrier state. *Science* 80(230), 1157–1160 (1985).
- 55 Babinet C, Farza H, Morello D, Hadchouel M, Pourcel C. Specific expression of hepatitis B surface antigen (HBsAg) in transgenic mice. *Science* 230(4730), 1160–1163 (1985).
- 56 Ali BA, Huang TH, Salem HH, Xie QD. Expression of hepatitis B virus genes in early embryonic cells originated from hamster ova and human spermatozoa transfected with the complete viral genome. *Asian J. Androl.* 8(3), 273–279 (2006).
- 57 Aguayo F, Castillo A, Koriyama C *et al.* Human papillomavirus-16 is integrated in lung carcinomas: a study in Chile. *Brit. J. Cancer* England, 85–91 (2007).



- 58 Farza H, Hadchouel M, Scotto J, Tiollais P, Babinet C, Pourcel C. Replication and gene expression of hepatitis B virus in a transgenic mouse that contains the complete viral genome. *J. Virol.* 62(11), 4144–4152 (1988).
- 59 Yokoyama M, Nakao Y, Yang X *et al.* Alterations in physical state and expression of human papillomavirus type 18 DNA following crisis and establishment of immortalized ectocervical cells. *Virus Res.* 37, 139–151 (1995).
- 60 Dejean A, Carloni G, Brechot C, Tiollais P, Wain-Hobson S. Organization and expression of hepatitis B sequences cloned from hepatocellular carcinoma tissue DNA. *J. Cell. Biochem.* (3), 293–301 (1982).
- 61 Ikuta K, Ding M, Zhang F, Sixbey JW, Scott RS. Epithelial cell retention of transcriptionally active, P3HR-1-derived heterogeneous Epstein–Barr virus DNA with concurrent loss of parental virus. *J. Virol.* 85(15), 7634–7643 (2011).
- 62 Livezey KW, Negorev D, Simon D. Hepatitis B virus-transfected Hep G2 cells demonstrate genetic alterations and *de novo* viral integration in cells replicating HBV. *Mutat. Res.* 452(2), 163–178 (2000).
- 63 Müller K, Heller H, Doerfler W. Foreign DNA integration. Genome-wide perturbations of methylation and transcription in the recipient genomes. *J. Biol. Chem.* 276(17), 14271–14278 (2001).
- 64 Hochstein N, Muiznieks I, Mangel L, Brondke H, Doerfler W. Epigenetic status of an adenovirus type 12 transgenome upon long-term cultivation in hamster cells. *J. Virol.* 81(10), 5349–5361 (2007).
- 65 Prusty BK, Krohne G, Rudel T. Reactivation of chromosomally integrated human herpesvirus-6 by telomeric circle formation. *PLoS Genet.* 9(12), e1004033 (2013).
- 66 Endo A, Watanabe K, Ohye T *et al.* Molecular and virological evidence of viral activation from chromosomally integrated human herpesvirus 6A in a patient with X-linked severe combined immunodeficiency. *Clin. Infect. Dis.* 59(4), 545–548 (2014).
- 67 Gravel A, Dubuc I, Morissette G, Sedlak RH, Jerome KR, Flamand L. Inherited chromosomally integrated human herpesvirus 6 as a predisposing risk factor for the development of angina pectoris. *Proc. Natl Acad. Sci. USA* 112(26), 8058–8063 (2015).
- 68 Rouzine IM, Weinberger AD, Weinberger LS. An evolutionary role for HIV latency in enhancing viral transmission. *Cell* 160(5), 1002–1012 (2015).
- 69 Kaufer BB, Jarosinski KW, Osterrieder N. Herpesvirus telomeric repeats facilitate genomic integration into host telomeres and mobilization of viral DNA during reactivation. *J. Exp. Med.* 208(3), 605–615 (2011).
- 70 Kishi M, Harada H, Takahashi M *et al.* A repeat sequence, GGGTTA, is shared by DNA of human herpesvirus 6 and Marek's disease virus. *J. Virol.* 62(12), 4824–4827 (1988).
- 71 Secchiero P, Nicholas J, Deng H *et al.* Identification of human telomeric repeat motifs at the genome termini of human herpesvirus 7: structural analysis and heterogeneity. *J. Virol.* 69(12), 8041–8045 (1995).
- 72 Morissette G, Flamand L. Herpesviruses and chromosomal integration. *J. Virol.* 84(23), 12100–12109 (2010).
- 73 Gentile I, Borgia G. Vertical transmission of hepatitis B virus: challenges and solutions. *Int. J. Womens Health* 6(1), 605–611 (2014).
- 74 Garolla A, Pizzol D, Bertoldo A, Menegazzo M, Barzon L, Foresta C. Sperm viral infection and male infertility: focus on HBV, HCV, HIV, HPV, HSV, HCMV and AAV. *J. Reprod. Immunol.* 100(1), 20–29 (2013).
- 75 Raybould R, Fiander A, Wilkinson GWG, Hibbitts S. HPV integration detection in CaSki and SiHa using detection of integrated papillomavirus sequences and restriction-site PCR. *J. Virol. Methods* 206, 51–54 (2014).
- 76 Smith EM, Ritchie JM, Summersgill KF *et al.* Age, sexual behavior and human papillomavirus infection in oral cavity and oropharyngeal cancers. *Int. J. Cancer* 108(5), 766–772 (2004).
- 77 Akagi K, Li J, Broutian TR *et al.* Genome-wide analysis of HPV integration in human cancers reveals recurrent, focal genomic instability. *Genome Res.* 24(2), 185–199 (2014).
- 78 Dooley KE, Warburton A, McBride AA. Tandemly integrated HPV16 can form a brd4-dependent super-enhancer-like element that drives transcription of viral oncogenes. *MBio.* 7(5), 2–11 (2016).
- 79 Olthof NC, Speel E-JM, Kolligs J *et al.* Comprehensive analysis of HPV16 integration in OSCC reveals no significant impact of physical status on viral oncogene and virally disrupted human gene expression. *PLoS ONE* 9(2), e88718 (2014).
- 80 Das P, Thomas A, Kannan S *et al.* Human papillomavirus (HPV) genome status & cervical cancer outcome: a retrospective study. *Indian J. Med. Res.* 142(5), 525 (2015).
- 81 Matulic M, Soric J. Papillomavirus genomes in human cervical carcinoma: analysis of their integration and transcriptional activity. *Neoplasma* 41(2), 95–100 (1994).
- 82 Groves IJ, Knight ELA, Ang QY, Scarpini CG, Coleman N. HPV16 oncogene expression levels during early cervical carcinogenesis are determined by the balance of epigenetic chromatin modifications at the integrated virus genome. *Oncogene* 35, 1–14 (2016).
- 83 Forman D, de Martel C, Lacey CJ *et al.* Global burden of human papillomavirus and related diseases. *Vaccine* 30(Suppl. 5), F12–F23 (2012).
- 84 Marongiu L, Godi A, Parry JV, Beddows S. Human papillomavirus 16, 18, 31 and 45 viral load, integration and methylation status stratified by cervical disease stage. *BMC Cancer* 14(1), 384 (2014).



- 85 Datta S, Chatterjee S, Veer V. Recent advances in molecular diagnostics of hepatitis B virus. *World J. Gastroenterol.* 20(40), 14615–14625 (2014).
- 86 Westrich JA, Warren CJ, Pyeon D. Evasion of host immune defenses by human papillomavirus. *Virus Res.* 231, 21–33 (2016).
- 87 Lace MJ, Anson JR, Haugen TH, Dierdorff JM, Turek LP. Interferon treatment of human keratinocytes harboring extrachromosomal, persistent HPV-16 plasmid genomes induces *de novo* viral integration. *Carcinogenesis* 36(1), 151–159 (2015).
- 88 Chen J-Y, Hsu H-C, Lee C-S, Chen D, Zuckerman A, Harrison T. Detection of hepatitis B virus DNA in hepatocellular carcinomas in Japan. *J. Virol. Methods* 19, 257–263 (1988).
- 89 Imazeki F, Yaginuma K, Omata M, Okuda K, Kobayashi M, Koike K. RNA transcripts of hepatitis B virus in hepatocellular carcinoma. *Hepatology* 7(4), 753–757 (1987).
- 90 Yaginuma K, Kobayashi H, Yoshida E, Kobayashi M, Koike K. Direct evidence for the expression of integrated hepatitis B virus DNA in human hepatoma tissue. *Gann* 75(9), 743–746 (1984).
- 91 Zhang X-K, Egan J, Huang D, Sun Z, Chien V, Chiu J-F. Hepatitis B virus DNA integration and expression of an ERB B-like gene in human hepatocellular carcinoma. *Biochem. Biophys. Res. Commun.* 188(1), 344–351 (1992).
- 92 Bowyer SM, Dusheiko GM, Schoub BD, Kew MC. Expression of the hepatitis B virus genome in chronic hepatitis B carriers and patients with hepatocellular carcinoma. *Proc. Natl Acad. Sci. USA* 84(3), 847–850 (1987).
- 93 Chen WN, Oon CJ, Leong AL, Koh S, Teng SW. Expression of integrated hepatitis B virus X variants in human hepatocellular carcinomas and its significance. *Biochem. Biophys. Res. Commun.* 276(3), 885–892 (2000).
- 94 Saigo K, Yoshida K, Ikeda R *et al.* Integration of hepatitis B virus DNA into the myeloid/lymphoid or mixed-lineage leukemia (*MLL4*) gene and rearrangements of *MLL4* in human hepatocellular carcinoma. *Hum. Mutat.* 29(5), 703–708 (2008).
- 95 Lu X, Lin Q, Lin M *et al.* Multiple-integrations of *HPV16* genome and altered transcription of viral oncogenes and cellular genes are associated with the development of cervical cancer. *PLoS ONE* 9(7), e97588 (2014).
- 96 Nakanishi G, Fujii K, Asagoe K, Tanaka T, Iwatsuki K. Human papillomavirus genome integration in multifocal vulvar Bowen's disease and squamous cell carcinoma. *Clin. Exp. Dermatol.* 34(8), e965–e967 (2009).
- 97 Aguayo F, Khan N, Koriyama C *et al.* Human papillomavirus and Epstein–Barr virus infections in breast cancer from Chile. *Infect. Agent Cancer* 6(1), 7 (2011).
- 98 Deng Z, Hasegawa M, Kiyuna A, Suzuki M. Viral load, physical status and E6/E7 mRNA expression of human papillomavirus in head and neck squamous cell carcinoma. *Head Neck* 35(6), 800–808 (2013).
- 99 Ho CM, Lee BH, Chang SF *et al.* Integration of human papillomavirus correlates with high levels of viral oncogene transcripts in cervical carcinogenesis. *Virus Res.* 161(2), 124–130 (2011).
- 100 Lehn H, Villa LL, Marziona F, Hilgarth M, Hillemans HG, Sauer G. Physical state and biological activity of human papillomavirus genomes in precancerous lesions of the female genital tract. *J. Gen. Virol.* 69(Pt 1), 187–196 (1988).
- 101 Lehn H, Krieg P, Sauer G. Papillomavirus genomes in human cervical carcinoma: analysis of their transcriptional activity. *Proc. Natl Acad. Sci. USA* 82, 5540–5544 (1985).
- 102 Hudelist G, Manavi M, Pischinger KID *et al.* Physical state and expression of HPV DNA in benign and dysplastic cervical tissue: different levels of viral integration are correlated with lesion grade. *Gynecol. Oncol.* 92(3), 873–880 (2004).
- 103 Hashida Y, Imajoh M, Nemoto Y *et al.* Detection of Merkel cell polyomavirus with a tumor-specific signature in non-small-cell lung cancer. *Br. J. Cancer* 108(3), 629–637 (2013).
- 104 Leitz M, Stieler K, Grundhoff A, Moll I, Brandner J. Merkel cell polyomavirus detection in Merkel cell cancer tumors in northern Germany using PCR and protein expression. *J. Med. Microbiol.* 86, 1813–1819 (2014).
- 105 Tsurimoto T, Fujiyama A, Matsubara K. Stable expression and replication of hepatitis B virus genome in an integrated state in a human hepatoma cell line transfected with the cloned viral DNA. *Proc. Natl Acad. Sci. USA* 84(2), 444–448 (1987).
- 106 Asselbergs FA, Will H, Wingfield P, Hirschi M. A recombinant Chinese hamster ovary cell line containing a 300-fold amplified tetramer of the hepatitis B genome together with a double selection marker expresses high levels of viral protein. *J. Mol. Biol.* 189(3), 401–411 (1986).
- 107 Fellig Y, Almog G, Galun E, Ketzin Gilad M. A hepatocellular carcinoma cell line producing mature hepatitis B viral particles. *Biochem. Biophys. Res. Commun.* 321(2), 269–274 (2004).
- 108 Takeshima H, Namiki M, Inokoshi J *et al.* Stable expression of hepatitis B virus genome in a primate kidney cell. *Arch. Virol.* 109, 35–49 (1989).
- 109 Jackson R, Rosa BA, Lameiras S *et al.* Functional variants of human papillomavirus type 16 demonstrate host genome integration and transcriptional alterations corresponding to their unique cancer epidemiology. *BMC Genomics* 17(1), 851 (2016).
- 110 Chen RW, Aalto Y, Teesalu T *et al.* Establishment and characterization of human papillomavirus type 16 DNA immortalised human tonsillar epithelial cell lines. *Eur. J. Cancer* 39(5), 698–707 (2003).
- 111 Kaur P, McDougall JK. Characterization of primary human keratinocytes transformed by human papillomavirus type 18. *J. Virol.* 62(6), 1917–1924 (1988).

- 112 Mitrani-Rosenbaum S, Kitron N. Integration and transcription of human papillomavirus type 6 recombinant DNA in mouse cells. *Virus Res.* 8(4), 335–347 (1987).
- 113 Bagis H, Arat S, Mercan HAO, Cirakoglu B. Stable transmission and expression of the Hepatitis B virus total genome in hybrid transgenic mice until F10 generation. *J. Exp. Zool.* 305, 420–427 (2006).

